UNIWERSYTET WARSZAWSKI

Wydział Biologii
Instytut Genetyki i Biotechnologii
Prof. dr hab Joanna Kufel

Warszawa, 12-08-2019

Recenzja pracy doktorskiej mgr Piotra Jana Machtela pt. "Opracowanie metody identyfikacji ryboprzełączników bakteryjnych z wykorzystaniem wysokoprzepustowego sekwencjonowania (Term-SHAPE-seq) oraz zestawu metod nacelowanych na analizę ich struktury i aktywności transkrypcyjnej"

Przedstawiona do recenzji praca doktorska mgr Piotra Jana Machtela została wykonana pod kierunkiem dr hab. Kamilli Bąkowskiej-Żywickiej w Zespole Transkryptomiki Funkcjonalnej Instytutu Chemii Bioorganicznej PAN w Poznaniu. Mgr Machtel jest pierwszym autorem pracy przeglądowej w *J. Appl. Genet.* (2016), współautorem publikacji w *FEBS Open Bio.* (2016) oraz pierwszym współautorem manuskryptu (obecnie w bazie *BioRxiv* i w trakcie recenzji w *Sci. Rep.*).

Prawidłowe funkcjonowanie wszystkich organizmów, także jednokomórkowych, wymaga ścisłej koordynacji ekspresji genów, szczególnie podczas adaptacji do zmieniających się warunków. Odbywa się to głównie przez regulację transkrypcji, wydajności dojrzewania i stabilności mRNA, przebiegu translacji oraz poziomu, właściwości (modyfikacje) i stabilności białek. W komórkach bakteryjnych jeden z mechanizmów regulacji polega na wykorzystaniu ryboprzełączników, które poprzez wiązanie swoistych ligandów mają wpływ na strukturę cząsteczek mRNA, co umożliwia specyficzne oddziaływanie z szeregiem czynników białkowych lub regulatorowych cząsteczek ncRNA. W swojej pracy doktorskiej mgr Machtel podjął się opracowania wyskoprzepustowego podejścia identyfikacji ryboprzełączników transkrypcyjnych, które opiera się na połączeniu dwóch znanych już wcześniej technik, Term-seq i SHAPE-seq. W rezultacie nowa metoda, adekwatnie nazwana Term-SHAPE-seq, pozwala na wyznaczenie miejsc przedwczesnej terminacji transkrypcji oraz struktury drugorzędowej RNA tworzącej ryboprzełącznik. Główne osiągnięcia wynikające z przeprowadzonych badań polegały przede wszystkim na opracowaniu, zoptymalizowaniu i zastosowaniu Term-SHAPE-seq do wyszukania i zbadania właściwości ryboprzełączników transkrypcyjnych u *Bacillus subtilis.* W rezultacie wykryto i zweryfikowano istnienie nowego ryboprzełącznika w sekwencji 5' UTR genu *yfmG* oraz określono jego parametry biochemiczne i

strukturalne, wykorzystując podejścia przetestowane wcześniej w pracy dla lepiej poznanego ryboprzełącznika SAM *metE*. Ciekawe są również obserwacje, że ryboprzełączniki tego samego typu (w tym przypadku klasy SAM-I) regulujące ekspresję różnych genów czy operonów cechują się różnym profilem aktywności na skutek braku ligandu (metioniny) oraz że poziom indukcji nie przekłada się bezpośrednio na zmianę ekspresji regulowanego genu/operonu.

W pracy zastosowano zarówno standardowe metody biologii molekularnej i biochemii dotyczące analizy kwasów nukleinowych, jak i specjalistyczne i zaawansowane techniki, również wysokoprzepustowego sekwencjonowania, służące do badań oddziaływania cząsteczek RNA-ligand oraz struktury drugorzędowej RNA *in vitro* i *in vivo* (SHAPE). Użycie tych technik było możliwe dzięki ekspertyzie dostępnej w laboratorium Kamilli Bąkowskiej-Żywickiej oraz w Instytucie Chemii Bioorganicznej PAN. Praca zawiera wszystkie niezbędne części i została przygotowana w sposób umożliwiający zrozumienie omawianych zagadnień i wykonanych badań. Rozprawa jest napisana po polsku, poprawnie, w sposób jasny i przejrzysty, bogatym i nawet czasem ozdobnym językiem, a błędy czy niezręczności językowe są nieliczne. Dysertacja jest starannie opracowana i ujęta pod względem struktury i zawartości, a także edytorskim, ale mogłaby być bardziej zwięzła.

W obszernym i usystematyzowanym wstępie autor wyczerpująco przedstawił tematykę związaną z regulacją ekspresji genów w organizmach bakteryjnych, czasem nawet nadmiernie szczegółowo i na zbyt podstawowym poziomie (np. treść początkowych rozdziałów 2.1.1-2.1.3, łącznie z przedstawieniem działania operonu *lacZYA*). Z kolei dogłębne omówienie terminacji transkrypcji bakteryjnej polimerazy RNA oraz regulatorowych funkcji cząsteczek RNA, w szczególności ryboprzełączników, i ich właściwości, poparte informatywnymi i obficie ilustrowanymi przykładami regulacji jest bardzo przydatne i przyczynia się do lepszego zrozumienia problematyki i zakresu badań zawartych w pracy. Bardzo sympatyczna jest także krótka opisowa historia dziejów RNA w rozdziale 2.2.1. Z perspektywy osoby zajmującej się zbliżoną tematyką w komórkach eukariotycznych brakuje może zwięzłego podkreślenia specyficzności omawianych w pracy typów regulacji dla komórek bakteryjnych. Niektóre elementy omawianych mechanizmów (np. przełączniki typu TTP) są obecne u eukariota, ale są zdecydowanie rzadziej spotykane, jest to co prawda omówione w rozdziale 2.2.4.2, ale przydałoby się bardziej stanowcze uwydatnienie tych różnic.

Cele pracy są jasno i precyzyjnie sformułowane i dokładnie określają zakres planowanych doświadczeń, aczkolwiek są ujęte zbyt obszernie i opisowo. Przede wszystkim przedstawienie technicznych szczegółów nie jest właściwe w rozdziale dotyczącym celów pracy. Część zawartości tego rozdziału zdecydowanie lepiej pasuje do sekcji wyników, jako uzasadnienie i wyjaśnienie konkretnych etapów pracy doświadczalnej. Materiały i metody są zaprezentowane poprawnie i kompetentnie, jednak wiele elementów technicznych zamiast w tej sekcji znajduje się w rozdziałach

poświęconych wynikom, brakuje także kilku istotnych informacji (wymienione w uwagach szczegółowych).

Część opisującą wyniki doświadczalne połączono z ich dyskusją, co jest uzasadnione w przypadku pracy o charakterze metodologicznym i składającej się w kilku zadań. Wyniki są przedstawione niezwykle skrupulatnie, ze szczególnym uwzględnieniem konkretnych procedur i z wyczerpującymi wyjaśnieniami ułatwiającymi prześledzenie kolejnych etapów. Potrzeba dokładnego ujęcia strony technicznej pracy poświęconej opracowaniu nowej metody jest zrozumiała, jednak jak na potrzeby dysertacji doktorskiej jest to potraktowane w sposób zbyt opisowy, brakuje rzeczowości i pewnej dyscypliny naukowej, na skutek czego praca jest nadmiernie rozbudowana i rozwlekła. Wiele szczegółów eksperymentalnych powinno być raczej umieszczone w części Materiały i Metody, a nie przy omawianiu wyników.

Na podstawie rezultatów pierwszego sekwencjonowania Term-SHAPE-Seq autor wnioskuje, że powiodło się wykrycie zachodzenia przedwczesnej terminacji transkrypcji, co może świadczyć o regulacji ekspresji z udziałem ryboprzełączników transkrypcyjnych. Założenia metody zostały potwierdzone przez identyfikację znanych struktur tego typu. Pytanie, ile z ryboprzełączników scharakteryzowanych dotychczas u *B. subtilis* zostało wykrytych przy pomocy Term-SHAPE-Seq, nie znalazłam takiej informacji w pracy. Ponadto, dzięki zastosowaniu modyfikacji RNA z użyciem NAI udało się uzyskać informacje dotyczące struktury badanych cząsteczek RNA. Za poprawnością tych danych przemawiają wygenerowane na ich podstawie struktury drugorzędowe $tRNA^{Thr}$ i 5S RNA. Ciekawe jak zgodność ta wygląda dla cząsteczek mRNA o mniej stabilnych i zdefiniowanych a bardziej zmiennych strukturach? Nie jestem pewna czy takie dane są dostępne dla mRNA *B. subtilis*, ale struktury niektórych mRNA zapewne były badane i warto pokusić się o porównanie. W przypadku *E. coli* analiza wysokoprzepustowa struktur RNA przy użyciu SHAPE ujawniła wiele ciekawych informacji dla cząsteczek mRNA (Mustoe et al., *Cell*, 2018). Wyszukiwanie nowych potencjalnych ryboprzełączników zostało omówione na dwóch przykładach (*srlX* w operonie *asnH* i w rejonie 5'UTR genu *yfmG*). Czy tylko te dwie sekwencje zostały zidentyfikowane jako potencjalne ryboprzełączniki, to dość skromna liczba, wydaje się, że odpowiednia analiza danych powinna wyłonić znacznie więcej takich przypadków, mimo zastosowania sekwencjonowania o średniej przepustowości. Mylące są sprzeczne stwierdzenia dotyczące *srlX*, pierwsze to „Pierwszym kandydatem na ryboprzełącznik jest sekwencja liderowa *srlX*" (str. 77), a drugie „(...) zdaje się być mało prawdopodobne by region *srlX* pełnił funkcję ryboprzełącznika". Proszę o komentarz w tej kwestii. Obecność ryboprzełącznika w rejonie 5'UTR genu *yfmG* została zweryfikowana pod kątem strukturalnym i funkcjonalnym z zastosowaniem dwóch z trzech wcześniej przetestowanych metod, jednak nie udało się ustalić natury ligandu wiążącego tę strukturę. Czy autor ma jakieś przypuszczenia co do możliwych ligandów innych niż testowane?

Z kolei wnikliwa analiza danych dotyczących indukcji ryboprzełączników SAM-I

doprowadziła do konkluzji, że profil indukcji jest swoisty dla każdego z genów/operonów, mimo że regulacja jest oparta o ten sam typ ryboprzełącznika. Okazało się, że dynamika indukcji jest powiązana z fizjologicznym znaczeniem tych genów w metabolizmie metioniny, siarki i pokrewnych szlaków, czyli najsilniejsze zmiany zachodzą dla czynników bezpośrednio zaangażowanych w odpowiedź na brak metioniny. Czy widoczna jest korelacja między poznanymi strukturami ryboprzełącznika obecnymi w każdym z genów, siłą oddziaływania z ligandem i stopniem indukcji? Czy jest możliwe przełożenie współczynnika IR na zmiany syntezy poszczególnych białek, np. na podstawie dostępnej literatury?

Ostatnia część pracy dotyczy modyfikacji usprawniającej analizę aktywności transkrypcyjnej ryboprzełączników poprzez wprowadzenie dodatkowego etapu cięcia transkryptów przy udziale RNazy H. Optymalizacja reakcji w warunkach *in vitro* opisana szczegółowo w rozdziale 4.4.1 jest tak naprawdę zbędna, ponieważ jest to bardzo dobrze opisane podejście, ponadto zastosowanie tego etapu dla całkowitego RNA wyizolowanego z bakterii stwarza zupełnie inną sytuację, co także zauważa sam autor. Szkoda, że nie zastosowano etapu denaturacji podczas procedury cięcia RNazą H w warunkach *ex vivo*, co zapewne zwiększyłoby wydajność reakcji. W przeciwieństwie do zbędnych elementów technicznych w poprzednich rozdziałach Wyników, opis metody ddPCR jest jak najbardziej na miejscu, ponieważ jest to podejście mniej znane i rzadziej wykorzystywane. Zastosowanie tej techniki jest w pełni właściwe, ponieważ pozwoliło na określenie bezwzględnego poziomu badanych transkryptów z dużą czułością. Czy próbowano więc wyznaczyć tą metodą stężenie cząsteczek mRNA *yxjH* i *yoaDCB*, których ilość była zbyt niska do detekcji metodą RT-qPCR? Mogłoby to pokazać, czy faktycznie ddPCR jest bardziej czuły i nadaje się do detekcji transkryptów o słabej ekspresji. Porównanie współczynników indukcji uzyskanych obiema metodami (RT-qPCR i ddPCR) pokazało, że wyniki nie są do końca spójne, czy autor mógłby jednoznacznie ocenić, która metoda jest bardziej wiarygodna? Ponadto, dziwi trochę wykonanie jednego powtórzenia biologicznego eksperymentów z użyciem ddPCR (wnioskuję to na podstawie braku analizy statystycznej).

Dyskusja uzyskanych wyników po każdym z kolejnych zadań w osobnych rozdziałach jest wnikliwa i rzeczowa. Zawiera przede wszystkim podsumowanie rezultatów i skupia się na technicznej stronie eksperymentów, co jest zrozumiałe w pracy o charakterze metodologicznym. Jednak w zbyt dużym stopniu pokrywa się z sekcją wyników i brakuje trochę bardziej generalnych wniosków. Końcowe podsumowanie pozwala na bardziej całościowe spojrzenie na osiągnięcia uzyskane w pracy, chociaż powinno być bardziej zwięzłe i spójne, a w mniejszym stopniu powielać informacje znajdujące się w streszczeniu i wynikach. Ciekawą propozycją są plany zastosowania nowego podejścia do zbadania regulacji przez ryboprzełączniki u bakterii o znaczeniu klinicznym, czyli *Staphylococcus aureus* opornej na metycyklinę.

Podsumowując, autor wykonał solidną pracę doświadczalną, eksperymenty zostały przeprowadzone w sposób zaplanowany i skrupulatny, aczkolwiek nie udało się uniknąć pewnych błędów i uproszczeń. Warto podkreślić, że przeprowadzone badania wymagają dużej systematyczności, dokładności i sprawności eksperymentalnej. Przede wszystkim autor wykazał, że opracowana przez niego metoda Term-SHAPE-Seq może służyć do wykrywania ryboprzełączników transkrypcyjnych i innych strukturalnych elementów regulatorowych oraz do badania dynamiki ich działania. Zaprezentowane w pracy wyniki zastosowania Term-SHAPE-Seq nie są w pełni satysfakcjonujące, bo powiodła się w rezultacie identyfikacja i weryfikacja tylko jednego funkcjonalnego ryboprzełącznika, co jednak świadczy o potencjale tego podejścia. Autor logicznie wyjaśnił przyczyny takiego stanu rzeczy i zaproponował przebieg dalszego postępowania. Jak sam sugeruje, konieczna jest dalsza optymalizacja i udoskonalenie tego podejścia aby było bardziej wydajne i wiarygodne.

Dysertacja pana mgr Machtela jest na wysokim poziomie i świadczy o jego samodzielności i dojrzałości naukowej. Praca spełnia ustawowe wymagania stawiane rozprawom doktorskim, wnioskuję zatem do Rady Instytutu Chemii Bioorganicznej PAN o dopuszczenie pana mgr Piotra Jana Machtela do dalszych etapów przewodu doktorskiego.

Uwagi szczegółowe:

1. Wstęp, str 21, co oznacza „polimeraza RNA zaczyna oddziaływać z RNAP"?, to prawdopodobnie przejęzyczenie.

3. Eksperyment w rozdziale 4.1.2. wykonano przy pomocy półilościowych i ilościowych reakcji RT-PCR, a współczynnik indukcji ryboprzełączników został policzony na podstawie różnicy poziomu formy FL i T. Czy transkrypty te, choćby niektóre z nich, można wykryć techniką Northern blot, co pozwoliłoby na bezpośrednią detekcje obu form? Czy jest możliwe zastosowanie innej metody wykrywania i przeliczania ilości tych transkryptów. Ograniczenia spowodowanego różnicowym sposobem określania stężenia cząsteczek nie niweluje zastosowane później w pracy udoskonalenia polegające na cięciu mRNA przez RNazę H i zastosowaniu tachniki ddPCR.

4. Rys 13. Warto byłoby przedstawić, które wyniki są istotne statystycznie.

5. Str 72. Nukleofil nie jest odczynnikiem tylko grupą chemiczną która jest donorem elektronu; w przypadku reakcji acylacji RNA przez NAI nukleofilem (jak to zwykle bywa dla RNA) jest grupa 2'OH w rybozie, a NAI ma właściwości elektrofilityczne.

6. Str 72. Jakość próbek RNA przed przygotowaniem bibliotek powinna być sprawdzona przy użyciu urządzenia Bioanalizer.

7. Str 74. Nie jest dobrze wyjaśnione na czym polegają charakterystyczne dla ryboprzełączników sygnatury świadczące o zachodzeniu mechanizmu przedwczesnej terminacji. Oczywiście można się łatwo domyśleć o co chodzi, ale przydałoby się jaśniejsze objaśnienie.

8. Str 76. Czy struktura drugorzędowa 5S RNA jest zgodna z opublikowaną, nie jest to wspomniane.

9. Rys 23, 25, 28A, 32A, 33A, 40 nie są konieczne w pracy doktorskiej.

10. Rys 44. W tym eksperymencie należało przeprowadzić dodatkową kontrolę polegającą na przeprowadzeniu całej reakcji bez obecności RNazy H, jest to niezbyt ważna kontrola, ale warto byłoby ją zrobić dla jednego zestawu starterów.

11. Braki w opisie niektórych protokołów, szczególnie jeśli były przeprowadzane „zgodnie z protokołem dostarczonym przez producenta", często jednak wprowadza się pewne modyfikacje. Czy reakcję odwrotnej transkrypcji podczas przygotowania bibliotek wykonano tak samo jak w przypadku analizy RNA *in vitro*? Nie jest również podane w jakiej temperaturze przeprowadzano reakcję hybrydyzacji oligomeru DNA przed cięciem RNazą H. Jest to akurat dość istotne ponieważ nie udało się uzyskać cięcia w warunkach *ex vivo* dla aż siedmiu transkryptów. Większość protokołów cięcia RNazą H uwzględnia etap denaturacji RNA, co znacznie zwiększa wydajność hybrydyzacji oligomeru, a co za tym idzie także cięcia RNA. Autor sam przyznaje, że nie zastosował tego etapu.

12. Dlaczego w Tabelach 5, 6, 7 i 8 oraz na Rys 61 podano wartości IR* bez SD lub SE? Czy doświadczenie było wykonane tylko w jednym powtórzeniu biologicznym?

13. Drobne pomyłki, nieliczne błędy językowe, składniowe czy literówki („termsensory", „ateunacja", „asreus", „krzywa wzrostu *B. subtilis*, hodowanej"; zbyt częste stosowanie zwrotu „dlatego też"; używanie terminów anglojęzycznych, np. *spacer* zamiast łącznik, 3' linker zamiast linker 3')

14. Spis skrótów jest bardziej praktyczny na początku pracy. Niektóre skróty nie są objaśnione (np. DFHBI).

Prof. dr hab. Joanna Kufel

UNIWERSYTET WARSZAWSKI
Instytut Genetyki i Biotechnologii
Prof. Joanna Kufel